

# Trends and Practices in Process Capability Studies

Mahendra Saha \*

Department of Statistics

Central University of Rajasthan

Bandarsindri, 305 817, India

Sudhansu S. Maiti

Department of Statistics

Visva Bharati University

Santiniketan, 731 235, India

## Abstract

Quantifying the “capability” of a manufacturing process is an important initial step in any quality improvement program. Capability is usually defined in dictionaries as “the ability to carry out a task, to achieve an objective”. Process capability indices(PCIs) is defined as a combination of materials, methods, equipments and people engaged in producing a measurable output. PCIs which establish the relationships between the actual process performance and the manufacturing specifications, have been a focus of research in quality assurance and process capability analysis. Capability indices that qualify process potential and process performance are practical tools for successful quality improvement activities and quality program implementation. As a matter of fact, all processes have inherent statistical variability, which can be identified, evaluated and reduced by statistical methods. Generalized Process Capability Index, defined as the ratio of proportion of specification conformance (or, process yield) to proportion of desired (or, natural) conformance. We review the process capability indices in case of normal, non-normal, discrete and multivariate process distributions and discuss the inferential aspects of some of these process capability indices. Relations among the process capability indices have also been illustrated with examples. Finally we also consider the process capability indices using conditional ordering and transforming multivariate data to univariate one using the concept of structural function.

---

\*Corresponding author. e-mail: mahendrasaha@curaj.ac.in

# 1 Introduction

Process capability is an important concept for industrial managers to understand. The challenge in today's competitive markets is to be on the leading edge of producing high quality products at minimum costs. This cannot be done without a systematic approach and this approach is contained within what has been called "Statistical quality control". The segment of statistical quality control(SQC) discussed here is the process capability study. Process capability indices(PCIs) aim to quantify the capability of a process of quality characteristic to meet some specifications that are related to a measurable characteristic of its produced items. These specifications are determined through the lower specification limit ( $L$ ), the upper specification limit ( $U$ ) and the target value ( $T$ ). A variety of such indices have been developed in last two decades. Following the work of Kane (1986), a number of articles have appeared introducing new indices and /or studying the properties of existing ones. Kaminsky et al. (1998) have given their critical comments on uses of these indices, and suggested a future measurement. Uses of process capability indices in administration and supplier certification have been discussed in Latzko (1985) and Schneider et al. (1995). Excellent reviews on them are given by Rodriguez (1992), Kotz and Johnson (1993, 2002), and Kotz and Lovelace (1998). In addition, Spiring et al. (2003) provide an extensive bibliography on process capability indices. It appears to be a general acceptance of the idea that PCIs can be used only after it has been established that a process is in "statistical control" (for example, by the use of control charts). This is reasonable if it simply requires that there be no irregular changes in quality level. However, there seems to be, in some quarters, an assumption that the measured characteristic should have a normal distribution (at least approximately), although it is difficult to see why a good industrial process must result in a normal distribution for every measured characteristic. The majority of the process capability indices discussed in the literature are associated only with processes that can be described through some continuous distributions of the characteristics and, in particular, normally distributed characteristics. The most widely used such indices are  $C_p$  [by Juran (1974)],  $C_{pk}$  [by Kane (1986)],  $C_{pm}$  [by Hsiang and Taguchi (1985)] and  $C_{pmk}$  [by Choi and Owen (1990), Pearn et al. (1992)] and their generalizations for non-normal processes suggested by Clements (1989), Constable and Hobbs (1992), Pearn and Kotz (1994-1995), Pearn and Chen (1995) and Mukherjee (1995). A number of new approaches to process capability analysis have been attempted by Carr (1991) and Flaig (1996). Another index that is given by

Boyles (1994) is also worth mentioning. Often, however, one is faced with processes described by a characteristic whose values are discrete. Therefore, in such cases none of these indices can be used. The indices suggested so far whose assessment is meaningful regardless of whether the studied process is discrete or continuous are those suggested by Yeh and Bhattacharya (1998), Borges and Ho (2001), Perakis and Xekalaki (2002, 2005) and Maiti et al. (2010).

In this article, our objective is to look into the overview of successive development of the process capability indices from the very beginning and also the inferential aspects of these capabilities, not only in case of normal process distribution, but also for non-normal and in some cases for discrete process distributions. The article has been organized as follows. Section 2 deals with some notations, where as Section 3 consider the background of process capability indices which have been developed in the earlier stage. Section 4, 5 and 6 consider the same in case of non-normal process distribution, any general process distribution, whether the process is normal, non-normal or discrete process distribution and multivariate process distributions respectively. Section 6 is devoted to gives concluding remarks.

## 2 Notations

- $L$  : Lower specification limit.
- $U$  : Upper specification limit.
- $f(x)$  : Probability density (or mass) function of quality characteristic  $X$ .
- $F(x)$  : Cumulative distribution function of quality characteristic  $X$ .
- $\bar{F}(x)$  :  $1 - F(x)$ .
- $UDL$  : Upper desired limit.
- $LDL$  : Lower desired limit.
- $\alpha_1$  :  $P(X < LDL)$ .
- $\alpha_2$  :  $P(X > UDL)$ .
- $\mu$  : Mean of the distribution.
- $\mu_e$  : Median of the distribution.
- $\sigma$  : Standard deviation of the distribution.
- $p$  : Process yield, *i.e.*,  $\int_L^U f(x)dx = F(U) - F(L)$ .
- $p_o$  : Desirable yield,  $F(UDL) - F(LDL) = 1 - \alpha$  with  $\alpha = \alpha_1 + \alpha_2$ .

### 3 Background

Process Capability analysis is a technique that has application in many segments of the product cycle, including product and process design, vendor sourcing, production or manufacturing planning and manufacturing. Quality of the products has always been a major concern for both consumers and producers. There appears to be a general acceptance of the idea that PCIs can be used only after it has been established that a process is in “statistical control”. This is reasonable, if it simply required that there be no irregular changes in quality level. However, there seems to be, in some quarters, an assumption that the measured characteristic should have a normal distribution, although it is difficult to see why a good industrial process must result in a normal distribution for every measured characteristic. The first PCI was developed by Juran(1974) and is defined as

$$\begin{aligned} C_p &= \frac{U - L}{6\sigma} \\ &= \frac{d}{3\sigma}, \end{aligned} \tag{3.1}$$

where,  $d = (U - L)/2$ . Note that  $C_p$  does not depend on process mean. A  $C_p$  value of 1 means that 99.73% of all individual items will be within specification. However, if there is only a slight change in the process mean or a slight increase in the process variation, a larger portion of items will be out of specification. Currently,  $C_p$  values of 1.33 or greater are used widely as industry standards. Some companies now require  $C_p$  values of 1.66 from their suppliers. A  $C_p$  of 1.33 indicates that 99.73% of products production is within  $(100\%/1.33) = 75\%$  of the specification, whereas a  $C_p$  of 1.66 shows that almost all product measurements fall within  $(100\%/1.66) = 60\%$  of the specification. For example, if the width of the specification is about 20 and the standard deviation is 3,

$$C_p = \frac{30 - 10}{18} = 1.1$$

indicating that not only can the process produce almost all production within specifications but also  $100\%/1.11 = 90\%$  of the specification will contain essentially all the production measurements. Whereas  $C_p$  values represent the setting in which the process is centered between the specification limits, process are frequently operated with the mean other than halfway between the upper and lower specification limits.

When the process mean is off-centered of the specification, the result is that one specification

limit (the closer to the process mean) becomes the focal point of the process capability calculation. The modified capability index, represents the situation in which one specification limit is more likely to be exceeded. The  $C_{pk}$  is then introduced by Kane(1986) to reflect the impact of  $\mu$  (process mean) on the process capability indices.

$$\begin{aligned} C_{pk} &= \min \left[ \frac{U - \mu}{3\sigma}, \frac{\mu - L}{3\sigma} \right] \\ &= \left[ \frac{d - |\mu - M|}{3\sigma} \right], \end{aligned} \quad (3.2)$$

where,  $M = (U + L)/2$ . The change in the denominator from six to three standard deviations is the result of the two one-sided quality concerns. For example, the  $C_{pk}$  for the situation described above is  $C_{pk} = \min[0.78, 1.44] = 0.78$ . When  $C_p$  and  $C_{pk}$  are calculated for a process that is centered, the two calculation methods give the same number. Therefore,  $C_{pk}$  is generally preferred because it is not dependent on the process being centered. However, the disadvantage of using only the  $C_{pk}$  value is that it does not reveal whether a poor process capability is the result of a process having a large dispersion or an off-centered process.

Therefore, both  $C_p$  and  $C_{pk}$  should be used to judge the process capability. If the index values are different, then this is a sign that the process mean is off-center and considerable improvement in the proportion of acceptable product can be made by centering the process mean.  $C_{pk}$  has been criticized because it does not measure the process deviation from target. A process may have a high  $C_{pk}$  even when the process mean is off-target and close to the specification limits as long as the process spread is small.

As an alternative, process capability index,  $C_{pm}$ , given by Hsiang and Taguchi(1985), takes into account the influence of the departure of the process mean  $\mu$  from the target  $T$ ,

$$\begin{aligned} C_{pm} &= \frac{U - L}{6\sqrt{\sigma^2 + (\mu - T)^2}} \\ &= \frac{d}{3\sqrt{\sigma^2 + (\mu - T)^2}} \\ &= \frac{d}{3\sqrt{E(X - T)^2}} \end{aligned} \quad (3.3)$$

where  $\mu$  is the process mean and  $T$  is the target value and  $E(.)$  denotes “expected value”. It is easy to verify that  $C_{pm}$  will possess necessary properties required for assessing process capability. If the process variance( i.e.,  $\sigma^2$  ) increases(decreases) the  $C_{pm}$  will decrease(increase). If the process drifts from its target value (i.e., if  $\mu$  moves away from  $T$ ) then  $C_{pm}$  decreases. In

the case where both the process variance and the process mean change, the  $C_{pm}$  index reflects these changes as well. Usually,  $T = M$  ; if  $T \neq M$  the situation is sometimes described as “asymmetric tolerances” (see Boyles (1994) and Vannman (1997b, 1998a)). The measure  $C_{pm}$  sometimes called the “Taguchi index”. There is also the hybrid index given by Pearn et al. (1992)

$$\begin{aligned} C_{pmk} &= \min \left[ \frac{U - \mu}{3\sqrt{\sigma^2 + (\mu - T)^2}}, \frac{\mu - L}{3\sqrt{\sigma^2 + (\mu - T)^2}} \right] \\ &= \frac{d - |\mu - M|}{3\sqrt{E(\mu - T)^2}} \end{aligned} \quad (3.4)$$

Clearly,  $C_{pmk}$  is only meaningful when deviation from target is the main concern. This index is based on the quadratic loss function and, thus, should only be used when there is evidence of a quadratic monetary loss. There are an infinite number of possible loss functions, and in many practical cases, material outside specification limits may result in a total loss rather than a quadratic loss. Clearly  $C_p \geq C_{pk} \geq C_{pmk}$  and  $C_p \geq C_{pm} \geq C_{pmk}$ . The relation between  $C_{pk}$  and  $C_{pm}$  is less clear cut. From Equations (1) and (2) we have

Boyles (1994) suggested index is

$$S_{pk} = \frac{1}{3}\Phi^{-1} \left[ \frac{1}{2}\Phi \left( \frac{U - \mu}{\sigma} \right) + \frac{1}{2}\Phi \left( \frac{\mu - L}{\sigma} \right) \right] \quad [Boyles, (1994)] \quad (3.5)$$

An enlightening view of relations among our basic PCIs can be obtained from studies of the superstructure PCIs” introduced by Vannman (1995). Let  $d = (U - L)/2$  be the half length of the specification interval. Vannman(1995) defined a class of PCIs which is a generalization of the existing indices. It depends on two non-negative parameters  $u$  and  $v$

$$C_p(u, v) = \left[ \frac{d - u}{6\sqrt{\sigma^2 + v(\mu - T)^2}} \right] ; \quad [Vannman, 1995] \quad (3.6)$$

where  $\mu$  is the process mean,  $\sigma$  is the process standard deviation,  $T$  is the target value and  $M = (U + L)/2$  is the mid-point of the specification interval. By varying the values of  $u$  and  $v$ , one can easily see that  $C_p(0, 0) = C_p$  (Juran (1974)),  $C_p(1, 0) = C_{pk}$  (Kane (1986)),  $C_p(0, 1) = C_{pm}$  (Hsiang and Taguchi (1985)), and  $C_p(1, 1) = C_{pmk}$  (Pearn, Kotz and Johnson(1992)). Moreover, when  $T = M$  and for  $u, v > 0$ ,  $C_p(u, v)$  has the same interpretation as  $C_{pm}$  and  $C_{pmk}$ .

Spiring (1997) also defines a PCI,  $C_p^{(w)} = C_p(0, w)$ . However, in this definition,  $w$  is not necessarily a constant; it may be a function of  $\frac{\mu-T}{\sigma}$ . In principle this allows  $[w(\mu - T)/\sigma]^2$  in Vannman's formula to be replaced by any function of  $(\mu - T)/\sigma$ . So, in effect,

$$C_p^{(w)} = \frac{C_p}{1 + g((\mu - T)/\sigma)^2} ; \quad [Spiring, 1997] \quad (3.7)$$

with a general choice of function  $g(\cdot)$ , though for practical purposes it should be a positive, increasing function.

## 4 PCIs under the assumption of non-normal process distribution

As already noted, the “6” in Equation (3.1) has been associated with the idea that a normal distribution for  $X$  provides a satisfactory approximation. Of course, both practitioners and theoreticians realized that this would not always be the case, and some (at least) of the second group energetically busied themselves with the task of coming up with relevant information and suggestions. Some practitioners, on the other hand, have claimed that  $C_p$  need not be assessed on the grounds of direct relevance to properties of NC product, though it is not clear what other means of assessment are to be used.

At a relatively early date, Clements (1989), in an influential paper, suggested that “ $6\sigma$ ” be replaced by the length of the interval between the upper and lower 0.135 percentage points of the distribution of  $X$  (this is for a normal  $N(\mu, \sigma^2)$ ). The new PCI is

$$C'_p = \frac{U - L}{\xi_{1-a} - \xi_a} \quad [Clements, (1989)] \quad (4.8)$$

where  $\xi_a$  is defined by  $P[X \leq \xi_a] = a$ , taking  $a = 0.00135$ , so that  $\xi_{1-a}$ ,  $\xi_a$  are the upper and lower 0.135 percentiles of the distribution of  $X$ . [For a  $N(\mu, \sigma^2)$  distribution  $\xi_{1-a} = \mu + 3\sigma$ ,  $\xi_a = \mu - 3\sigma$ .]

Clements (1989) suggested fitting a Pearson system distribution for  $X$ , in order to obtain the required  $\xi_a$  values. Applications of this kind of method, with various assumed distributional forms, have been quite numerous since 1992. References include: Rodriguez (1992), Bittanti et al. (1998), Lovera et al. (1997) -all Pearson system; Castagliola (1996) -Burr distributions; Farnum (1996), Polansky et al. (1998), Pyzdek (1992) -all Johnson system; Padgett and Sengupta (1996) -Weibull and log-normal; Mukherjee and Singh (1997) -Weibull; Sarkar and Pal (1997)

-extreme value; Somerville and Montgomery (1996/7) -t, gamma, and log-normal; Sundaraiyar (1996) -inverse Gaussian. As mentioned above, Polansky (1998, 2001) uses a general kernel fitting method.

An Index defined on Inter quantile range is given by

$$I = \frac{U - L}{F^{-1}(1 - \alpha_2) - F^{-1}(\alpha_1)} \quad [Mukherjee, (1995)] \quad (4.9)$$

Yeh and Bhattacharya (1998) proposed use of a PCI based on the ratios of expected proportion NC to actual observed or estimated proportion NC. Let  $X$  be the random variable associated with the characteristic of a process under study. Here we assume that  $X$  is positive and has a continuous distribution. Let  $p_0$  be the desired proportion of non-conforming output and  $p$  be the actual proportion of non-conforming. Then

$$a) \quad C_p = \frac{p_0}{p} ; \quad [Yeh and Bhattacharya, 1998] \quad (4.10)$$

Yeh and Chen (1999) have extended this to multivariate cases. Another PCI, suggested by Yeh and Bhattacharya (1998), distinguishes between NC items for which  $X$  is less than  $L$ , and those for which  $X$  is greater than  $U$ . Let  $\alpha_0^L$  and  $\alpha_0^U$  be the expected proportions of non-conforming products the manufacturer can tolerate on the lower and upper specification limits, respectively. Let  $\alpha_L = P(X < L)$  and  $\alpha_U = P(X > U)$  be the actual proportions of non-conformance of the process. The question of whether the process is “capable” can then be answered by comparing  $\alpha_0^L$  to  $\alpha_L$  and  $\alpha_0^U$  to  $\alpha_U$ . Thus

$$b) \quad C_f = \min \left[ \frac{\alpha_0^L}{\alpha_L}, \frac{\alpha_0^U}{\alpha_U} \right] ; \quad [Yeh and Bhattacharya, 1998] \quad (4.11)$$

Here the index is directly linked to the probability of non-conformance. The computation involved in estimating this index is more intensive than the conventional PCIs. This perhaps is not a serious constraints in today’s computer technology. A new measure of process capability which is in one-to-one correspondence with the process non-conforming fraction  $\pi$ . Namely

$$C = \frac{1}{3} \Phi^{-1} \left( 1 - \frac{\pi}{2} \right) \quad [Borges and Ho, (2001)] \quad (4.12)$$

where  $\pi = 1 - P(L \leq X \leq U)$  is the in-control process fraction defective. The index  $C$  has some interesting properties. (i) It is one-to-one correspondence with the in-control process fraction defective  $\pi$ . Processes with the same fraction defective will be equally capable. Moreover, the



process capability will respond only to changes in the process fraction defective and not to changes in the distribution of the observed quality characteristic.

(ii) If, under process stability,  $X$  is normally distributed with mean  $\mu = (L+U)/2$  and standard deviation  $\sigma$ , then

$$\frac{\pi}{2} = \Phi\left(\frac{L-\mu}{\sigma}\right) = 1 - \Phi\left(\frac{U-\mu}{\sigma}\right)$$

and

$$C_p = \frac{U-\mu}{3\sigma} = \frac{1}{3}\Phi^{-1}(1 - \pi/2) = C,$$

i.e., process capability is measured by the  $C_p$  value of a centered normal process with the same in-control process fraction defective  $\pi$ . This property makes  $C$  values easy to interpret and understand. (iii) It can be applied to discrete and continuous, uni- and multivariate quality indicators and no assumption on the distribution of the observed quality characteristic is made.

## 5 PCIs under the assumption of any general process distribution

Now a PCI is introduced which overcomes the drawbacks of the standard indices discussed earlier. It is defined as the ratio

$$C_{pc} = \frac{1-p_0}{1-p}; \quad [Perakis \text{ and } Xekalaki, 2002] \quad (5.13)$$

Where  $p_0$  denotes the minimum allowable proportion of conformance.  $p$  is the process yield lying between  $L$  and  $U$ . The value of  $p_0$  must be intuitively close to unity and depends on the nature of the examined process and the requirements of the customers. This index is used for process with unilateral as well as bilateral tolerances and can take into account the minimum acceptable process yield(if such quantity has been set).

A generalized process capability index given by Maiti et al. (2010), which is the ratio of proportion of specification conformance (or, process yield) to proportion of desired (or, natural) conformance. Almost all previously defined process capabilities are directly or indirectly associated with this generalized one.

$$\begin{aligned} C_{py} &= \frac{F(U) - F(L)}{F(UDL) - F(LDL)} \\ &= \frac{p}{p_0}. \end{aligned} \quad (5.14)$$

The following theorem gives different forms of  $C_{py}$  when  $X$  has uniform distribution.

**Theorem 5.1** *If  $X$  follows uniform distribution over  $(a, b)$ ,  $a < b$ , then  $C_{py}$  reduces to*

- (a)  $C_p = \frac{U-L}{6\sigma}$  when  $LDL = \mu - 3\sigma$  and  $UDL = \mu + 3\sigma$  [Juran (1974)];
- (b)  $\dot{C}_p = \frac{U-L}{\xi_{1-\frac{\alpha}{2}} - \xi_{\frac{\alpha}{2}}}$  when  $LDL = \xi_{\frac{\alpha}{2}}$  and  $UDL = \xi_{1-\frac{\alpha}{2}}$  with  $\kappa = P(X \leq \xi_{\kappa})$  [Clements (1989)];
- (c)  $I = \frac{U-L}{F^{-1}(1-\alpha_2) - F^{-1}(\alpha_1)}$  when  $LDL = F^{-1}(\alpha_1)$  and  $UDL = F^{-1}(1 - \alpha_2)$  [Mukherjee (1995)].

When proportions of desired conformances are  $\alpha_1$  for lower tail and  $1 - \alpha_2$  for upper tail, then  $C_{py} = \frac{p}{1-\alpha_1-\alpha_2}$ , which for a normal process with  $LDL = \mu - 3\sigma$  and  $UDL = \mu + 3\sigma$  is  $C_{py} = \frac{p}{0.9973}$ .

Theorem 5.1 motivates us to formulate the generalized index  $C_{py}$  given in (5.14). Let us examine the behaviour of the proposed index  $C_{py}$  for different values of  $p$ . Obviously, if  $p = p_0$  (for normal distribution,  $p = 0.9973$ ), then  $C_{py} = 1$ . If the process yield  $p$  is greater than  $p_0$ , then  $C_{py} > 1$  whereas, if  $p < p_0$ , then  $C_{py} < 1$ , and the value of the index approaches zero as  $p$  tends to zero. Thus, the smallest possible value of  $C_{py}$  is zero.

Now the situation is of importance when the process is off-centered, *i.e.*,  $F(L) + F(U) \neq 1$ , but the proportion of desired conformance is achieved. In that case they defined the index as follows:

$$\begin{aligned} C_{pyk} &= \min \left\{ \frac{F(U) - F(\mu_e)}{\frac{1}{2} - \alpha_2}, \frac{F(\mu_e) - F(L)}{\frac{1}{2} - \alpha_1} \right\} \\ &= \min \{C_{pyu}, C_{pyl}\} \end{aligned} \quad (5.15)$$

When  $C_{py}$  and  $C_{pyk}$  are calculated for a centered process, they come out to be the same. Therefore,  $C_{pyk}$  is generally preferred because it is not dependent on the process being centered. If the index values are different, then this is a sign that the process median (mean in case of symmetric distribution) is off-centered, and considerable improvement in the proportion of acceptable product can be made by centering the process median.

When  $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ , then

$$\begin{aligned} C_{pyk} &= \min \left\{ \frac{F(U) - \frac{1}{2}}{\frac{1}{2}(1 - \alpha)}, \frac{\frac{1}{2} - F(L)}{\frac{1}{2}(1 - \alpha)} \right\} \\ &= \frac{d - |\frac{1}{2} - F(M)|}{\frac{1}{2}(1 - \alpha)}, \end{aligned}$$

where  $F(M) = \frac{F(L)+F(U)}{2}$ .

For a normal process and under the assumption of uniformity as in Theorem 5.1,  $C_{pyk}$  reduces to

$$\begin{aligned} C_{pk} &= \min \left\{ \frac{U - \mu}{3\sigma}, \frac{\mu - L}{3\sigma} \right\} \\ &= \frac{d - |\mu - M|}{3\sigma}. \end{aligned}$$

where  $d = \frac{U-L}{2}$ .

It generally happens that process target  $T$  is such that  $F(T) = \frac{F(L)+F(U)}{2}$ ; if  $F(T) \neq \frac{F(L)+F(U)}{2}$ , the situation may be described as “generalized asymmetric tolerances” [Boyles (1994) and Vannman (1997,1998) have described by the term “asymmetric tolerances” when  $T \neq M = \frac{L+U}{2}$ ]. Under this circumstance, they defined

$$C_{pTk} = \min \left\{ \frac{F(U) - F(T)}{\frac{1}{2} - \alpha_2}, \frac{F(T) - F(L)}{\frac{1}{2} - \alpha_1} \right\}. \quad (5.16)$$

At this point, it would be interesting to note that  $C_{pTk}$  is equal to  $C_{pyk}$  when  $T = \mu_e$ .

This generalized process capability index overcomes many deficiencies of the PCIs which have been already discussed and more or less all the existing PCIs are directly or indirectly associated with this index. And it can be used for process with unilateral or bilateral tolerances.

## 6 PCIs under the assumption of multivariate process distribution

A more precise title for this Section would be “PCIs for Use When  $X$  is Multivariate”. Many of the PCIs in this group are not, in fact, multivariate. May be they should be, but writers have opted for construction of univariate PCIs, based on the multivariate distributions of  $X$ . Nevertheless we will term them all “multivariate PCIs” (MPCIs). References with a title including the word “multivariate” or “bivariate” are: Beck and Ester (1998); Bernardo and Irony (1996); Boyles (1996b); Chan et al. (1991); Davis et al. (1992); Hellmich and Wolf (1996); Hubele et al. (1991); Karl et al. (1994); Li and Lin (1996); Mukherjee and Singh (1994); Niverthi and Dey (2000); Shariari et al. (1995); Taam et al. (1993); Tang and Barnett (1998); Veevers (1995, 1998, 1999); Wang et al. (2000); Wierda (1992, 1993, 1994a, 1994b); and Yeh and Chen (1999). Multivariate situations are also discussed in the following references,

that do not indicate, explicitly, in their titles that this is so: Chan et al. (1988b); Wang and Chen (1998/9); and Wang and Hubele (1999, 2001).

The univariate specification interval ( $L \leq X \leq U$ ) is now replaced by a specification region. This may just be constructed from separate specification intervals: one for each variable  $X_i$  in  $X$ . The specification region is then the hyperrectangle

$$\prod_{i=1}^v (L_i \leq X_i \leq U_i).$$

However, more complex regions may be used, reflecting perceived relations among the variables in  $X$ . These are of the general form

$$L \leq g(X) \leq U.$$

Often,  $L$  is zero. Possibly for mathematical convenience,  $g(X)$  is often taken as a monotonic function of the joint probability density function of  $X$ . Thus if  $X$  is assumed to have a multivariate normal  $N_v(\mu, \Sigma)$  distribution, one might take

$$g(X) = (X - \mu)' \Sigma^{-1} (X - \mu) \tag{6.17}$$

and regard an item as NC if  $g(X) > U$ . In this way we obtain the ellipsoidal specification region

$$(X - \mu)' \Sigma^{-1} (X - \mu) \leq U. \tag{6.18}$$

An analogue of  $C_p$  is

$$\frac{\text{Volume of}\{(X - \mu)' \Sigma^{-1} (X - \mu) \leq U\}}{\text{Volume of}\{(X - \mu)' \Sigma^{-1} (X - \mu) \leq R\}} = \left(\frac{U}{R}\right)^v, \tag{6.19}$$

where

$$P[(X - \mu)' \Sigma^{-1} (X - \mu) \leq R] = 1 - p.$$

If the distribution of  $X$  is multivariate normal then  $(X - \mu)' \Sigma^{-1} (X - \mu)$  has a  $\chi^2$  distribution with  $v$  degrees of freedom and  $R = \chi_{v, 1-p}^2$  (the upper 100(1 -  $p$ )% point of the  $\chi^2$  or “chi-squared” distribution with  $v$  degrees of freedom).

Chen (1994) applies this method to the case when the specification region is of the form in Equation (6.18). The region is defined by

$$\max_{i=1,2,\dots,v} \left( \frac{X_i - M_i}{d_i} \right) \leq 1$$

with  $M_i = (L_i + U_i)/2$  and  $d_i = (U_i - L_i)/2$ , and Chen defined  $MC_P$  as  $R^{-1}$ , where

$$P \left[ \max_{i=1,2,\dots,v} \left( \frac{X_i - M_i}{d_i} \right) \leq R \right] = 1 - p. \quad (6.20)$$

There can be many variants on these approaches. For example, the  $g(X)$  in Equation (6.17) might be replaced by  $(X - \mu)'A^{-1}(X - \mu)$  where  $A$  is a positive matrix, not necessarily the variance-covariance matrix of the distribution of  $X$ .

Shariari et al. (1995) proposed a truly multivariate MPCl. It contains three components. The first is of the type in Equation (6.20). The second is the significance level of the Hotelling's  $T^2$  statistic

$$T^2 = n(\bar{X} - \mu)'S^{-1}(\bar{X} - \mu),$$

which is

$$P \left[ F_{v,n-v} > \frac{n-v}{v(n-1)} T^2 \right],$$

where  $F_{v,n-v}$  denotes a variable having the  $F$  distribution with  $v, n-v$  degrees of freedom. The final component just takes values 1 or 0 according to a modified process region -defined as the smallest region similar in shape to the specification region, circumscribed about a specified probability contour (of the distribution of  $X$ ) -is or not entirely contained in the specification region.

Wang et al. (2000) compared this 3-component MPCl with Chen's (1994)  $MC_p$  and with an index  $MC_{pm}$  proposed by Taam et al. (1993) which is also a ratio of two volumes. The volume in the denominator is the same as in Equation (6.19) with  $R = \chi_{v,1-p}^2$  with  $p = 0.0027$ , while in the numerator we have the volume of a "modified specification region" which is the largest ellipsoid centered at the target that is within the original specification region.

Bairamov (2006) introduced the conditionally ordered order statistics for multivariate observations. Let  $X_1, X_2, \dots, X_n \in S \subseteq R^m$  be i.i.d. random vectors with  $m$ -variate c.d.f.  $F(x)$  and p.d.f.  $f(x)$ , where  $x = (x_1, x_2, \dots, x_m)$  and  $S$  is the support of  $X$ . Consider the real-valued function  $N(X) : R^m \rightarrow R, x = (x_1, x_2, \dots, x_m)$ , which is continuous in its arguments satisfying  $N(X) \geq 0$  for all  $x \in R^m$  with  $N(X) = 0$  if and only if  $x = 0$ , where  $0 = (0, 0, \dots, 0)$ .  $N(X_1), N(X_2), \dots, N(X_n)$  are i.i.d. random variables with c.d.f.  $P(N(X_i) \leq N(x)), x \in R$ . The function  $N(x)$  introduces partial ordering among the random vectors  $X_1, X_2, \dots, X_n$  and  $X_1$  is said to be conditionally less than  $X_2$  (or  $X_1$  precedes  $X_2$ ) if  $N(X_1) \leq N(X_2)$ . This ordering is

denoted by  $X_1 \prec X_2$ .

Selection of the function  $N(x)$  is independent of the structure of the system and it can be chosen in various ways depending on the conditions available in the process/system. The following selections for  $N(x)$  may be of special interest.

- (a)  $N(x_1, x_2, \dots, x_n) = \sum_{i=1}^n a_i x_i$ .
- (b)  $N(x_1, x_2, \dots, x_n) = \min(x_1, x_2, \dots, x_n)$ .
- (c)  $N(x_1, x_2, \dots, x_n) = \max(x_1, x_2, \dots, x_n)$ .

While the first function reflects the combined effects with the weights  $a_i$ , the last two functions consider the extreme values.

For multivariate quality characteristic, if there are lower and upper specification vectors  $\underline{L}$  and  $\underline{U}$  given, then  $\underline{L} \prec \underline{X} \prec \underline{U}$  is more meaningful and similarly for lower desirable vector  $\underline{LDL}$  (or lower tolerance vector  $\underline{LTL}$ ) and upper desirable vector  $\underline{UDL}$  (or upper tolerance vector  $\underline{UTL}$ ). In that circumstance Maiti et al. (2013) defined generalized process capability index for multivariate process quality characteristic as

$$\begin{aligned}
C_{py}^M &= \frac{P(\underline{L} \prec \underline{X} \prec \underline{U})}{P(\underline{LDL} \prec \underline{X} \prec \underline{UDL})} \\
&= \frac{P(N(\underline{L}) < N(\underline{X}) < N(\underline{U}))}{P(N(\underline{LDL}) < N(\underline{X}) < N(\underline{UDL}))} \\
&= \frac{F(N(\underline{U})) - F(N(\underline{L}))}{F(N(\underline{UDL})) - F(N(\underline{LDL}))} \\
&= \frac{p}{p_0}
\end{aligned} \tag{6.21}$$

In the similar fashion as in the univariate case they defined  $C_{pyk}^M$  and  $C_{pTk}^M$  as

$$C_{pyk}^M = \min \left\{ \frac{F(N(\underline{U})) - \frac{1}{2}}{\frac{1}{2} - \alpha_2}, \frac{\frac{1}{2} - F(N(\underline{L}))}{\frac{1}{2} - \alpha_1} \right\}$$

for off-centered process and

$$C_{pTk}^M = \min \left\{ \frac{F(N(\underline{U})) - F(N(\underline{T}))}{\frac{1}{2} - \alpha_2}, \frac{F(N(\underline{T})) - F(N(\underline{L}))}{\frac{1}{2} - \alpha_1} \right\}$$

for off-target process respectively, where  $F(\cdot)$  is the cumulative distribution function of  $N(\cdot)$ .

They described the steps to be followed to calculate process capability index for a set of multivariate quality characteristic data.

1. Suppose  $\underline{X}_1, \underline{X}_2, \dots, \underline{X}_n$  are  $n$  observation vectors come from a assumed multivariate distribution.

2. Transform the observations using  $N(\underline{X})$ . Let  $Y_1 = N(\underline{X}_1), Y_2 = N(\underline{X}_2), \dots, Y_n = N(\underline{X}_n)$ . Here the distribution function  $F(y) = P(N(\underline{X}) < y)$  is the structural function associated with  $\underline{X}$  via  $Y = N(\underline{X})$ . Now  $N(\underline{X})$  will possess a distribution.
3. Now assuming  $Y_1, Y_2, \dots, Y_n$  are observations, we will find a best fitted distribution for the observations using Q-Q plot technique or any goodness of fit test.
4. Using this best fitted univariate distribution, PCI is to be calculated following the approach of Maiti et al. (2010).
5. If there are more than one multivariate distributions assumed and corresponding structural functions are considered, then we will choose that structural function for which the transformed data set is best fitted. Here, to define generalized process capability index for multivariate data, conditional ordering has been used, and multivariate data has been transformed to univariate one using the concept of structural function. After transforming the data, it is as simple as applying generalized process capability indices for univariate data given by Maiti et al. (2010). It can be well understood and comfortably be used by the practitioners.

## 7 Conclusions

Process capability studies play an important role in the process control since they assist to decide whether a manufacturing process is suitable and the applications meet the necessary quality standards. The assessment of process capability, which now appears to be quite simple, involves significant dimensions in a practical set up. This is due to the fact that some of the conditions necessary to establish process capability not fully satisfiable. These conditions stipulate that the process has to be under control, that the process output has to follow the normal distribution and the observed values of the quality characteristics be statistically independent. Since these conditions are not fully satisfied in many manufacturing situations, the process capability analysis is becoming to be a critical issue. This situation was of course realized by many researchers and resulted in numerous publications in the literature. Therefore, process capability analysis is valid only when the process under investigation is free of any special or assignable causes (i.e., is in-control.).

Recently, the research in the theory and practice of multivariate process capability indices has been very sparse in comparison to the research dealing with the univariate case. At present, for the multivariate capability indices, consistency in the methodology for evaluating this capa-

bility is not so much developed. Moreover, it is quite difficult to obtain the relevant statistical properties. Obviously, further investigations in this field are strongly desirable. In further correspondence, we will no doubt provide a fertile growth for new developments in the theory, methodology, statistical properties and applications of the PCIs.

## References

1. Bairamov, I. G. (2006): Progressive type II censored order statistics for multivariate observations, *Journal of Multivariate Analysis*, **97**, 797-809.
2. Beck, C. and Ester, S. (1998): Multicriteria Capability Indexes, *Spektrum*, **24**, 179-187.
3. Bernardo, J. M. and Irony, T. X. (1996): A General Multivariate Bayesian Process Capability Index, *Statistician*, **45**, 487-502.
4. Bittanti, S., Lovera, M. and Moiraghi, L. (1998): Application of Non-normal Process Capability Indices to Semiconductor Quality Control, *IEEE Transactions on Semiconductor Quality Control*, **11**, 296-302.
5. Boyles, R. A. (1994): Process Capability with Asymmetric Tolerances, *Communications in Statistics - Simulation and Computation*, **23**, 613-643.
6. Borges, W. and Ho, L. L. (2001): A Fraction Defective based Capability Index, *Quality and Reliability Engineering International*, **17**, 447-458.
7. Boyles, R. A. (1996b): Multivariate Process Analysis With Lattice Data, *Technometrics*, **38**, 37-49.
8. Carr, W. A. (1991): A new process capability index: Parts per million, *Quality Progress*, **24(8)**, 152.
9. Chan, L. K., Cheng, S. W. and Spiring, F. A. (1988b): A New Measure of Process Capability,  $C_{pm}$ , *Journal of Quality Technology*, **20**, 162-175.
10. Chen, H. (1994): A Multivariate Process Capability Index Over a Rectangular Solid Tolerance Zone, *Statistica Sinica*, **4**, 749-758.



11. Castagliola, P. (1996): Evaluation of Non-normal Process Capability Indices using Burr's Distribution, *Quality Engineering*, **8**, 587-593.
12. Choi, B. C. and Owen, D. B. (1990): A Study of a New Process Capability Index, *Communications in Statistics - Theory and Methods*, **19**, 1232-1245.
13. Clements, J. A. (1989): Process Capability Calculations for Non-normal Distributions, *Quality Progress*, **22**, 95-100.
14. Constable, G. K. and Hobbs, J. B. (1992): Small samples and non-normal capability, *ASQC Quality Congress Transactions - Nashville*, 37-43.
15. Davis, R. D., Kaminsky, F. C. and Saboo, S. (1992): Process Capability Analysis for Processes With Either a Circular or a Spherical Tolerance Zone, *Quality Engineering*, **5**, 41-54.
16. Farnum, N. R. (1996): Using Johnson Curves to Describe Non-normal Process Data, *Quality Engineering*, **9**, 329-336.
17. Flaig, J. J. (1996): A new approach to process capability analysis, *Quality Engineering*, **9**, 205-211.
18. Juran, J. M. (1974): Juran's Quality Control Handbook, 3rd ed. **McGraw-Hill**, New York, USA.
19. Hellmich, M. and Wolff, H. (1996): A New Approach for Describing and Controlling Process Capability for a Multivariate Process , *Proceedings of the ASA Section on Quality and Productivity*, 44-48.
20. Hsiang, T. C. and Taguchi, G. (1985): Tutorial on Quality Control and Assurance - The Taguchi Methods, *Joint Meeting of the American Statistical Association*, Las Vegas, Nevada, 188.
21. Hubele, N. F., Shahriari, H. and Cheng, C. S. (1991): A Bivariate Process Capability Vector, *Statistical Process Control in Manufacturing edited by J. B. Keats and D. C. Montgomery*, Marcel Dekker, New York, 299-310.

22. Kaminsky, F.C., Dovich, R.A. and Burke, R.J. (1998): Process Capability Indices ; Now and in the Future, *Quality Engineering*, **10**, 445-453.
23. Kane, V. E. (1986): Process Capability Indices, *Journal of Quality Technology*, **18**, 41-52.
24. Karl, D. P., Morisette, J. and Taams, W. (1994): Some Applications of a Multivariate Capability Index in Geometric Dimensioning and Tolerancing, *Quality Engineering*, **6**, 649-665.
25. Kotz, S. and Johnson, N. L. (1993): Process Capability Indices, **Chapman and Hall**, London, U.K.
26. Kotz, S. and Johnson, N. L. (2002): Process Capability Indices - A Review, *Journal of Quality Technology*, **34(1)**, 2-19.
27. Kotz, S. and Lovelace, C. (1998): Introduction to Process Capability Indices, **Arnold**, London, U.K.
28. Latzo, W. J. (1985): Process capability in administrative applications, *ASQC Quality Congress Transactions - Baltimore*, 168-173.
29. Li, Y. and Lin, C. (1996): Multivariate  $C_p$  Value, *Chinese Journal of Applied Probability and Statistics*, **12**, 132-138.
30. Lovera, C., Moiraghi, L., Montefusco, A. and Valentini, D. (1997): Estimation of Industrial Process Capabilities -Some Proposals and Studies to Overcome the Obstacle of Some Non-normal Distributions, *Abstract III, National Congress, SIMAI, Pavia, Italy*, 627-631.
31. Maiti, S. S., Saha, M., Nanda, A. K. (2010): On Generalizing Process Capability Indices, *Journal of Quality Technology and Quantitative Management* **7(3)**:279-300.
32. Mukherjee, S. P. (1995): Process Capability Indices and Associated Inference Problems, *Proceedings of the International Conference on tatistical Methods and Statistical Computation*, Seoul, South Korea, 243-249.
33. Mukherjee, S. P. and Singh, N. K. (1994): Sampling Properties of an Estimate of the Bivariate Process Capability Index, *Economic Quality Control*, **9**, 73-78.

34. Mukherjee, S. P. and Singh, N. K. (1997): Sampling Properties of an Estimator of a New Process Capability Index for Weibull Distributed Quality Characteristics, *Quality Engineering*, **10**, 291-294.
35. Niverthi, M. and Dey, D. K. (2000): Multivariate Process Capability - A Bayesian Perspective, *Communications in Statistics -Simulation and Computation*, **29**, 667-687.
36. Padgett, W. J. and Sengupta, A. (1996): Performance of Process Capability Indices for Weibull and Lognormal Distributions or Autoregressive Processes, *International Journal of Reliability, Quality and Safety Engineering*, **3**, 217-229.
37. Pearn, W. L. and Chen, K. S. (1995): Estimating Process Capability Indices for Non-normal Pearsonian Populations, *Quality and Reliability Engineering International*, **11**, 389-391.
38. Pearn, W. L. and Kotz, S. (1994-95): Application of Clements' Method for Calculating Second- and Third-generation Process Capability Indices for Non-Normal Pearsonian Populations, *Quality Engineering*, **7**, 139-145.
39. Pearn, W. L., Kotz, S. and Johnson, N. L. (1992): Distributional and Inferential Properties of Process Capability Indices, *Journal of Quality Technology*, **24**, 216-231.
40. Perakis, M. and Xekalaki, E. (2002): A Process Capability Index that is based on the Proportion of Conformance, *Journal of Statistical Computation and Simulation*, **72(9)**, 707-718.
41. Perakis, M. and Xekalaki, E. (2005): A Process Capability Index for Discrete Processes, *Journal of Statistical Computation and Simulation*, **75(3)**, 175-187.
42. Polansky, A. M. (1998): A Smooth Non-parametric Approach to Process Capability, *Quality and Reliability Engineering International*, **14**, 43-48.
43. Polansky, A. M., Chou, Y. -M. and Mason, R. L.(1998): Estimating Process Capability Indices for a Truncated Distribution, *Quality Engineering*, **11**, 257-265.
44. Polansky, A. M. (2001): A Smooth Non-parametric Approach to Multivariate Process Capability, *Technometrics*, **43**, 199-211.

45. Pyzdek, T. (1992): Process Capability Analysis Using Personal Computers, *Quality Engineering*, **4**, 419-440.
46. Rodriguez, R. N. (1992): Recent Developments in Process Capability Analysis, *Journal of Quality Technology*, **24**, 176-187.
47. Sarkar, A. and Pal, S. (1997): Estimation of Process Capability Index for Concentricity, *Quality Engineering*, **9**, 665-671.
48. Schneider, H., Pruett, J. and Lagrange, C. (1995): Uses of Process Capability Indices in Supplier Certification Process, *Quality Engineering*, **8**, 225-235.
49. Shahriari, H., Hubele, N. F. and Lawrence, F. P. (1995): A Multivariate Process Capability Vector, *Proceedings of the 4th Industrial Engineering Research Conference*, Nashville, TN, 303-308.
50. Spiring, F. A. (1997): A Unifying Approach to Process Capability Indices, *Journal of Quality Technology*, **29**, 49-58.
51. Spiring, F., Leung, B., Cheng, S. and Yeung, A. (2003): A Bibliography of Process Capability Papers, *Quality and Reliability Engineering International*, **19(5)**, 445-460.
52. Somerville, S. E. and Montgomery, D. C. (1996/7): Process Capability Indices and Non-normal Distributions, *Quality Engineering*, **9**, 305-316.
53. Sundaraiyar, V. H. (1996): Estimation of a Process Capability Index for Inverse Gaussian Distribution, *Communications in Statistics -Theory and Methods*, **25**, 2381-2398.
54. Taam, W., Subbaiah, P and Liddy, J. W. (1993): A Note on Multivariate Capability Indices, *Journal of Applied Statistics*, **20**, 339-351.
55. Tang, P. F. and Barnett, N. S. (1998): Capability Indices for Multivariate Processes, *Technical Report*, Division of Computation, Mathematics and Science, Victoria University, Melbourne, Australia.
56. Vannman, K. (1995): A Unified Approach to Capability Indices, *Statistica Sinica*, **5**, 805-820.

57. Vannman, K. (1997): A General class of Capability Indices in the case of Asymmetric Tolerances, *Communications in Statistics- Theory and Methods*, **26**, 2049-2072.
- Vannman, K. (1998): Families of Capability Indices for One-sided Specification Limits, *Statistics*, **31**, 43-66.
58. Veevers, A. (1995): A capability Index for Multiple Response, *CSIRO Mathematics and Statistics Report DMS 095*, Australia.
59. Veevers, A. (1998): Viability and Capability Indices for Multiresponse Processes, *Journal of Applied Statistics*, **25**, 545-558.
60. Veevers, A. (1999): "Capability indices for multiresponse processes" in statistical process monitoring and optimization edited by S. H. Park and G.G. Vining, Marcel-Decker, New York, NY, 241-256.
61. Wang, F. K., Hubele, N. F., Lawrence, F. P. and Miskulin, J. O. (2000): Comparison of Three Multivariate Process Capability Indices, *Journal of Quality Technology*, **32**, 263-275.
62. Wang, F. K. and Chen, J. C. (1998/9): Capability Index Using Principal Components Analysis, *Quality Engineering*, **11**, 21-27.
63. Wang, F. K. and Hubele, N. F. (1999): Quality Evaluation Using Geometric Distance Approach, *International Journal of Reliability, Quality and Safety Engineering*, **6**, 139-153.
64. Wang, F. K. and Hubele, N. F. (2001): Quality Evaluation of Geometric Tolerance Regions in Form and Location, *Quality Engineering*, **14**, 203-209.
65. Wierda, S. J. (1992): Multivariate Quality Control Estimation of the Percentage of Good Product, *Technical Report, University of Groningen, Groningen, Netherlands*.
66. Wierda, S. J. (1993): A Multivariate Process Capability Index, *Transactions of the ASQC Quality Congress*, 342-348.
67. Wierda, S. J. (1994a): Multivariate Statistical Process Control-Recent Results and Directions for Future Research, *Statistica Neerlandica*, **48**, 147-168.

- 68. Wierda, S. J. (1994b): Multivariate Statistical Process Control, Thesis, Wolters-Noordhof, Groningen, Netherlands.
- 69. Yeh, F. B. and Bhattacharya, S. (1998): A Robust Process Capability Index, *Communications in Statistics - Simulation and Computation*, **27(2)**, 565-589.
- 70. Yeh, A. B. and Chen, H. (1999): A Non-parametric Multivariate Process Capability Index, *Preprint*.